A Counter-based MAC Revisited: Towards Better Security

Eonkyung Lee[1]
Department of Applied Mathematics, Sejong University, Seoul, 143-747, Korea
eonkyung@sejong.ac.kr

Abstract. Bellare, Guérin, and Rogaway proposed a very efficient and secure message authentication scheme. By slightly modifying it, this article shows that the security is improved.

Keywords. Message authentication, Pseudorandom function, Adaptive chosen message attack.

1 Introduction

A message authentication scheme enables two parties sharing a key to authenticate their transmission. A *message authentication code* (MAC) is a special case of it, where the authenticated message is the sender's original message together with a tag. As security concerns grow, their role becomes more important. Bellare, Guérin, and Rogaway [1] constructed a counter-based MAC with noble features, especially better efficiency and security bound than a very popular MAC, the CBC MAC. These schemes are built on pseudorandom functions.

A *pseudorandom function* (PRF) is what looks like a truly random function (TRF). Since it is used as a primitive for building many cryptographic schemes, its security directly affects the schemes. Classically, security of a PRF is measured by distinguishing it from a TRF. This notion is denoted by PRF.

Desai and Miner [2] proposed an alternate notion to this, termed an *indistinguishable-uniform function* (IUF). They applied this notion to analyzing security of an arbitrary MAC viewing the whole of the MAC as a PRF primitive in the IUF sense.

Our result The aim of this article is twofold. On the one hand, we slightly change Bellare et al.'s MAC, and show that it has better security bound (i.e. its insecurity upper-bound is lower) than the original scheme in basic security as well as against replay attack. On the other hand, we concretely analyze the (computational) security of our MAC by using the notion of IUF in a different way from Desai et al.'s analysis for MAC, i.e. our scheme is not treated as a PRF itself but as a relatively complex protocol built on a PRF primitive. As a by-product, the security analysis is simplified not by evaluating explicitly the information-theoretic security of the MAC.

2 Preliminaries

Basic Notations $x \leftarrow S$ means that x is selected according to the probability distribution of S if S is a probability space, that x is selected uniformly at random from S if S is just a finite set, and that x is set to S if S is a value. For an integer $0 \leq x < 2^n$, $\langle x \rangle_n$ denotes the natural binary encoding of x as an n-bit string. For bit strings α and β, $\alpha.\beta$ denotes their concatenation.

[1] Every comment on this paper is welcome!

PRF: $\mathsf{Exp}_F^{\mathsf{PRF}}(A, b)$		IUF: $\mathsf{Exp}_F^{\mathsf{IUF}}(A, b)$	
1.	$a \leftarrow \mathsf{Keys}(F)$	1.	$a \leftarrow \mathsf{Keys}(F)$
2.	$\mathcal{O}_0 \leftarrow R_{l,L}$; $\mathcal{O}_1 \leftarrow F_a$	2.	$(x, s) \leftarrow A^{F_a}(\mathsf{find})$, where x was not a query.
3.	$d \leftarrow A^{\mathcal{O}_b}$	3.	$y_0 \leftarrow \{0,1\}^L$; $y_1 \leftarrow F_a(x)$
4.	Return d.	4.	$d \leftarrow A(\mathsf{guess}, y_b, s)$
		5.	Return d.

Figure 1: Experiments for security notions of PRFs

Pseudorandom Function Let $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a \in \mathsf{Keys}(F)}$ be a function family with a key space $\mathsf{Keys}(F)$, and let $R_{l,L}$ be the set of all functions from $\{0,1\}^l$ to $\{0,1\}^L$. Pseudorandomness of F can be measured in the notion of PRF or IUF. See Figure 1. Here A is an adversary, and it should be hard for A to distinguish between $b = 0$ and $b = 1$ in order for F to be secure. Insecurity of a PRF is defined as A's success probability.

Definition 1 Let F be a finite function family and let $\mathsf{N} \in \{\mathsf{PRF}, \mathsf{IUF}\}$. For any integers $t, q > 0$

$$\begin{aligned} \mathsf{Adv}_F^{\mathsf{N}}(A) &= \Pr[\mathsf{Exp}_F^{\mathsf{N}}(A,1) = 1] - \Pr[\mathsf{Exp}_F^{\mathsf{N}}(A,0) = 1], \\ \mathsf{Adv}_F^{\mathsf{N}}(t,q) &= \max_A\{\mathsf{Adv}_F^{\mathsf{N}}(A)\}, \end{aligned}$$

where the maximum is over all adversaries A with time complexity $\leq t$, making $\leq q$ oracle queries.

As follows, Desai et al. showed concrete relations between these notions, which proved tight.

Theorem 1 ([2]) *For any function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a \in \mathsf{Keys}(F)}$ and any integers $t, q > 0$*

$$\mathsf{Adv}_F^{\mathsf{IUF}}(t,q) \leq 2 \cdot \mathsf{Adv}_F^{\mathsf{PRF}}(t', q+1) \quad \textit{and} \quad \mathsf{Adv}_F^{\mathsf{PRF}}(t,q) \leq q \cdot \mathsf{Adv}_F^{\mathsf{IUF}}(t', q),$$

where $t' = t + O(l + L)$.

Proposition 2 ([2]) *There is a function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a \in \mathsf{Keys}(F)}$ such that*

$$\mathsf{Adv}_F^{\mathsf{PRF}}(t,q) \geq \frac{1}{2} \quad \textit{and} \quad \mathsf{Adv}_F^{\mathsf{IUF}}(t,q) \leq \frac{1}{q}$$

for any integers $t > 0$ and $0 < q \leq 2^{L-1}$.

3 Specification of MAC

The counter-based MAC, $\mathsf{MAC}' = (\mathsf{Sig}', \mathsf{Vf}')$, of Bellare et al. is built on a function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a \in \mathsf{Keys}(F)}$ and an integer $0 < b < l$, where it is written as $\mathsf{MAC}'_{F,b}$. Let $a \in \mathsf{Keys}(F)$ be the shared secret key between Sam and Vera. At the beginning of the protocol, Sam initializes his counter as zero for his signing algorithm Sig'. For any message M to be authenticated, it is supposed that its bit length is at most $b2^{l-b-1}$ and is a multiple of b (by standard padding arguments), and M is regarded as a sequence of b-bit blocks, i.e. $M = M[1] \dots M[n]$ where $M[i] \in \{0,1\}^b$ for all i. For this type of message M, a key $a \in \mathsf{Keys}(F)$, and a counter value $c \in \{0,1\}^{l-1}$, define

$$\mathsf{tag}_{F,b}(a, c, M) = F_a(0.c) \oplus F_a(1.\langle 1 \rangle_{l-b-1}.M[1]) \oplus \cdots \oplus F_a(1.\langle n \rangle_{l-b-1}.M[n]).$$

$\mathsf{Sig}'_{F,b}(a,C,M)$	$\mathsf{Vf}'_{F,b}(a,M',(C',z'))$	$\mathsf{Vf}_{F,b}(a,C,M',(C',z'))$
1. $C \leftarrow C+1$	1. $z \leftarrow \mathsf{tag}_{F,b}(a,\langle C'\rangle_{l-1},M')$	1. If $C' \neq C+1$, return 0 and stop.
2. $z \leftarrow \mathsf{tag}_{F,b}(a,\langle C\rangle_{l-1},M)$	2. If $z = z'$, return 1.	2. $z \leftarrow \mathsf{tag}_{F,b}(a,\langle C'\rangle_{l-1},M')$
3. Return (C,z).	Otherwise, return 0.	3. If $z = z'$, return 1 and $C \leftarrow C'$.
		Otherwise, return 0.

Figure 2: $\mathsf{MAC}'_{F,b} = (\mathsf{Sig}'_{F,b}, \mathsf{Vf}'_{F,b})$ and $\mathsf{MAC}_{F,b} = (\mathsf{Sig}_{F,b}, \mathsf{Vf}_{F,b})$, where $\mathsf{Sig}_{F,b} = \mathsf{Sig}'_{F,b}$.

Using this tagging function, $\mathsf{Sig}'_{F,b}(a,\cdot,\cdot)$ creates a signature given a message M and $\mathsf{Vf}'_{F,b}(a,\cdot,\cdot)$ verifies it. See Figure 2.

Since Vera's verification algorithm Vf' has no state, replay attack can be mounted. Namely, Vera accepts $(M,(C,z))$ sent by an adversary if Sam sent it before. To prevent this attack, the verification algorithm is made stateful. We will write this modified scheme as $\mathsf{MAC} = (\mathsf{Sig}, \mathsf{Vf})$, where the signing algorithm is the same as before (i.e. $\mathsf{Sig} = \mathsf{Sig}'$) and the verification algorithm is slightly changed so as to use a synchronized counter. See Figure 2.

How much does the modification weigh down the scheme? In usual environments, this change almost gives no additional burden to systems. The biggest extra load is that the counter of Vf should be initialized (as zero) at the same time as Sig. If DES or AES is used as F, the bit length of the counter is 63 or 127. In practice, before the initialization of it, the secret key a should be renewed. Therefore, the load will be negligible if Sam and Vera renew the key together with the counter initialization.

4 Notion of Security for MAC

A widely used notion of security for MACs is resistance to existential forgery under adaptive chosen message attack. Namely, for any adversary who is allowed to access an oracle for signing algorithm with his chosen messages, he hardly provides a new authenticated message. Here we can think an adversary who is intercepting and taking data travelling from signer to verifier. We will call this type of attack *intercepting and adaptive-chosen-message attack* ($\mathsf{IC}-\mathsf{CMA}$), and denote unforgeability under it by $\mathsf{UF}-\mathsf{IC}-\mathsf{CMA}$.

Case with stateful verifier. Next, let's consider another probable case. Adam gives a non-malicious message (e.g. "Would you lend me $100 for a while?") to Sam; Sam sends it (e.g. "I want to withdraw $100 from my account.") with his signature to a teller, Vera; Adam can watch this transmission. After some number of such transmission, Adam forges Sam's signature for a new message that is possibly malicious (e.g. "I want to transfer $1000 to Adam's account."), and then transmits the message together with the forged signature to Vera hoping to deceive her. Here, adversaries can only eavesdrop data travelling from signer to verifier. We will call this type of attack *eavesdropping and adaptive-chosen-message attack* ($\mathsf{ED}-\mathsf{CMA}$), and denote unforgeability under it by $\mathsf{UF}-\mathsf{ED}-\mathsf{CMA}$.

In many cases, it hardly affects unforgeability of a MAC whether the scheme is under $\mathsf{ED}-$ or $\mathsf{IC}-\mathsf{CMA}$ (i.e. whether or not verifier receives legitimately authenticated messages caused by adversary). However, in the case that verifier has a state, these attacks have different influences on security. To forge MAC under $\mathsf{ED}-\mathsf{CMA}$, the counter value for the forgery has to be consistent with

```
UF-ED-CMA: $\mathsf{Exp}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}_{F,b}}(A)$
1. $a \leftarrow \mathsf{Keys}(F)$; $C_0 \leftarrow 0$; $\mathcal{Q} \leftarrow \emptyset$
2. For $i = 1$ to $q$
    (a) $M_i \leftarrow A(\mathsf{query}, \mathcal{Q})$
    (b) $(C_i, z_i) \leftarrow \mathsf{Sig}_{F,b}(a, C_{i-1}, M_i)$, where $C_i = C_{i-1} + 1$
    (c) $\mathcal{Q} \leftarrow \mathcal{Q} \cup \{(M_i, (C_i, z_i))\}$
3. $(M, (C, z)) \leftarrow A(\mathsf{forge}, \mathcal{Q})$, where $M \notin \{M_1, \ldots, M_q\}$ and $C = C_q + 1$.
4. Output whatever $\mathsf{Vf}_{F,b}(a, C_q, M, (C, z))$ outputs.
```

Figure 3: Experiment for UF − ED − CMA of MAC with q-oracle queries

the previous values. So, experiment for UF − ED − CMA of MAC can be described as in Figure 3. We denote by $\mathsf{Adv}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}_{F,b}}(A)$ the probability that the outcome of the experiment $\mathsf{Exp}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}_{F,b}}(A)$ is 1. We associate to MAC its insecurity function, defined for any integers t, q, μ by

$$\mathsf{Adv}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}_{F,b}}(t, q, \mu) \quad = \quad \max_{A}\{\mathsf{Adv}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}_{F,b}}(A)\}.$$

Here the maximum is taken over all adversaries A with running time $\leq t$, number of oracle queries $\leq q$, and total message length (in oracle queries and in forgery) $\leq \mu$.

5 Security Analysis of MAC

Bellare et al. evaluated UF − IC − CMA of MAC′, and showed that their security bound is essentially tight.

Theorem 3 (Theorem 5 in [1]) *For any function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a \in \mathsf{Keys}(F)}$, and integers $0 < b < l$, $0 < q < 2^{l-1}$, and $t, \mu > 0$ where μ is a multiple of b*

$$\mathsf{Adv}^{\mathsf{UF-IC-CMA}}_{\mathsf{MAC}'_{F,b}}(t, q, \mu) \leq \mathsf{Adv}^{\mathsf{PRF}}_{F}(t', q') + 2^{-L}, \tag{1}$$

where $t' = t + O((l + L)q')$ and $q' = q + 1 + \mu/b$.

This computational security was shown using Theorem 4 in [1] for information-theoretic security.

5.1 UF − IC − CMA

Theorem 3 also holds for MAC. The reason is as follows. Under IC − CMA, signer is accessed many times during the query phase, and the signing algorithms of MAC and MAC′ are identical. So, adversaries come to get the same information from them each. On the other hand, the counter value for forgery of MAC′ has no significance in the proof of Theorem 3. It is chosen arbitrarily. So, the counter value for MAC can be chosen appropriately, i.e. the next value of verifier's counter. Therefore, MAC and MAC′ have the same security bound in the UF − IC − CMA sense.

5.2 UF − ED − CMA

For MAC′, the two notions UF − IC − CMA and UF − ED − CMA have no difference because its verifier is stateless. So we get the following, for which tightness argument for (1) also holds.

$$\mathsf{Adv}^{\mathsf{UF-ED-CMA}}_{\mathsf{MAC}'_{F,b}}(t, q, \mu) \leq \mathsf{Adv}^{\mathsf{PRF}}_{F}(t', q') + 2^{-L}. \tag{2}$$

For MAC, the argument is analogous to §5.1. The only difference is what the counter value for forgery of MAC is. It equals one plus the counter value for the last query in the eavesdropping case, while it is the counter value for the first query in the intercepting case. As discussed in §5.1, this causes no problem in the way to apply Theorem 3 here, and so we get

$$\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(t, q, \mu) \leq \mathsf{Adv}^{\mathsf{PRF}}_{F}(t', q') + 2^{-L}. \tag{3}$$

However, we evaluate $\mathsf{UF-ED-CMA}$ of MAC in a different way to tighten the above bound.

Theorem 4 *For any function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a\in\mathsf{Keys}(F)}$, and integers $0 < b < l$, $0 < q < 2^{l-1}$, and $t, \mu > 0$ where μ is a multiple of b*

$$\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(t, q, \mu) \leq \mathsf{Adv}^{\mathsf{IUF}}_{F}(t', q') + 2^{-L}, \tag{4}$$

where $t' = t + O((l + L)q')$ and $q' = q + \mu/b$.

Proof. For simplicity, let A be an adversary attacking $\mathsf{MAC}_{F,b}$ making q queries in time t such that the total length of all messages (in queries and in forgery) from it is μ-bits. We construct an adversary D for F as follows by running A and by accessing an oracle $\mathcal{O}$ which is F_a for some hidden key $a \leftarrow \mathsf{Keys}(F)$.

1. Run A: for $i = 1, \ldots, q$, let $M_i = M_i[1] \ldots M_i[n_i]$ be the i-th query from A; compute $z_i = \mathcal{O}(0.\langle i\rangle_{l-1}) \oplus \mathcal{O}(1.\langle 1\rangle_{l-b-1}.M_i[1]) \oplus \cdots \oplus \mathcal{O}(1.\langle n_i\rangle_{l-b-1}.M_i[n_i])$ by accessing $\mathcal{O}$; give (i, z_i) to A.
2. Let $(M, (C, z)) \leftarrow A$, where $C = q + 1$.
3. Decompose M into b-bit blocks: $M = M[1] \ldots M[n]$ for some proper n.
4. Compute $y_i = \mathcal{O}(1.\langle i\rangle_{l-b-1}.M[i])$ for all $1 \leq i \leq n$.
5. Let $x = 0.\langle C\rangle_{l-1}$ and $s = (z, y_1 \ldots y_n)$.
6. Output (x, s) as challenge, and receive y as the challenge response.
7. If $z = y \oplus y_1 \oplus \cdots \oplus y_n$, output 1. Otherwise, output 0.

In the above, D maintains the counter, incrementing it in the way of MAC, and implements $\mathsf{Sig}_{F,b}(a, \cdot, \cdot)$ given an oracle $\mathcal{O}$ for F_a. Note that it is possible that an $M[i]$ in the message M has been queried in the form of $1.\langle i\rangle_{l-b-1}.M[i]$ before, while C is a new value. Thus at the end of the find phase of D (at step 6), it should output not $M[i]$ but C, so that $x = 0.\langle C\rangle_{l-1}$. Since D cannot access $\mathcal{O}$ in its guess phase, it pre-computes $F_a(1.\langle i\rangle_{l-b-1}.M[i])$ for all i by accessing $\mathcal{O}$ during its find phase, and then passes them via state information s. In addition, z in the output of A is also passed via s in order to decide whether y is $F_a(x)$ or not.

Clearly, the total number of oracle queries made by D is $q' = q + \mu/b$, and the running time of D is $t' = t + O((l + L)q')$. The advantage of D is

$$\begin{aligned} \mathsf{Adv}^{\mathsf{IUF}}_{F}(D) &= \Pr[\mathsf{Exp}^{\mathsf{IUF}}_{F}(D, 1) = 1] - \Pr[\mathsf{Exp}^{\mathsf{IUF}}_{F}(D, 0) = 1] \\ &= \mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(A) - 2^{-L}. \end{aligned}$$

Since A is arbitrary,

$$\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(t, q, \mu) = \mathsf{Adv}^{\mathsf{IUF}}_{F}(D) + 2^{L} \leq \mathsf{Adv}^{\mathsf{IUF}}_{F}(t', q') + 2^{L}.$$

□

Our proof is much simpler than that of Theorem 3 (i.e. proofs of Theorems 4, 5 in [1]). A main reason is that ours does not explicitly evaluate the information-theoretic security, i.e. security of $\mathsf{MAC}_{F,b}$ when F is a TRF, due to well-meshing of IUF with $\mathsf{UF{-}ED{-}CMA}$.

How tight is our reduction? The following shows that the bound (4) is very tight.

Proposition 5 *There exist a function family $F = \{F_a : \{0,1\}^l \to \{0,1\}^L\}_{a\in\mathsf{Keys}(F)}$, and integers $0 < b < l$, $0 < q < 2^{l-1}$, and $t, \mu > 0$ such that*

$$\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(t, q, \mu) = \mathsf{Adv}^{\mathsf{IUF}}_{F}(t', q') + 2^{-L},$$

where μ is a multiple of b, $t' = t + O((l + L)q')$, and $q' = q + \mu/b$.

Proof. Choose any integers $0 < b < l$ and $0 < n \le 2^{l-b-1}$. For each $a \in \mathsf{Keys}(F)$, let F_a be any constant function. Then for any counter value C and any message $M = M[1]\dots M[n]$, $\mathsf{tag}_{F,b}(a, C, M) = 0\cdots0$ if n is odd and the value of F_a otherwise. Define an adversary A for $\mathsf{MAC}_{F,b}$ with oracle access to $\mathsf{Sig}_{F,b}(a, C_0, \cdot)$ as follows.

1. Select any $M_1 \in \{0,1\}^{bn}$, and get $(C_1, z_1) \leftarrow \mathsf{Sig}_{F,b}(a, C_0, M_1)$.
2. Select any $M \in \{0,1\}^{bn} - \{M_1\}$, and let $C \leftarrow C_1 + 1$ and $z \leftarrow z_1$.
3. Output $(M, (C, z))$.

Then $\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(A) = 1$, and so $\mathsf{Adv}^{\mathsf{UF\text{–}ED\text{–}CMA}}_{\mathsf{MAC}_{F,b}}(t, 1, 2bn) = 1$ where t is the running time of A.

On the other hand, the best adversary, say D, for F with oracle access to F_a will be

1. Select any $x' \in \{0,1\}^l$, and get $s \leftarrow F_a(x')$.
2. Select any $x \in \{0,1\}^l$ such that $x \neq x'$.
3. Output (x, s) as challenge, and receive y as the challenge response.
4. Output 1 if $y = s$, and 0 otherwise.

Since D queries once and runs in time $O(l + L)$, for $q' = 1 + 2n$ and $t' = t + O((l + L)q')$

$$\mathsf{Adv}^{\mathsf{IUF}}_{F}(t', q') = \mathsf{Adv}^{\mathsf{IUF}}_{F}(D) = \Pr[\mathsf{Exp}^{\mathsf{IUF}}_{F}(D, 1) = 1] - \Pr[\mathsf{Exp}^{\mathsf{IUF}}_{F}(D, 0) = 1] = 1 - 2^{-L},$$

from which the conclusion follows. □

5.3 Comparison of Securities of MAC and MAC′

In the sense of $\mathsf{UF{-}IC{-}CMA}$, MAC and MAC' have the same security bound as discussed in §5.1.

Let's consider the case of $\mathsf{UF{-}ED{-}CMA}$. Comparing MAC and MAC' directly, it is easy to see that MAC is at least as secure as MAC'. Let's look at their concrete securities. Theorem 1 and Proposition 2 imply that for any t and q, $\mathsf{Adv}^{\mathsf{IUF}}_{F}(t, q)$ is at most $2 \cdot \mathsf{Adv}^{\mathsf{PRF}}_{F}(t', q')$ and can be as small as $\frac{1}{q} \cdot \mathsf{Adv}^{\mathsf{PRF}}_{F}(t'', q'')$ for some t', q', q'', t''. From this fact, concrete security of MAC, (3) and (4), turns out to have better bound than MAC', (2).

References

[1] M. Bellare, J. Guérin, and P. Rogaway, *XOR MACs: New Methods for Message Authentication Using Finite Pseudoranom Functions*, CRYPTO '95, LNCS 963 (1995) 15–28

[2] A. Desai and S. Miner, *Concrete Security Characterizations of PRFs and PRPs: Reductions and Applications*, ASIACRYPT '00, LNCS 1976 (2000) 503–516